When does the top homology of a random simplicial complex vanish?

L. Aronshtam[*]   N. Linial[†]

February 27, 2013

Abstract

Several years ago Linial and Meshulam [7] introduced a model called $X_d(n, p)$ of random n-vertex d-dimensional simplicial complexes. The following question suggests itself very naturally: What is the threshold probability $p = p(n)$ at which the d-dimensional homology of such a random d-complex is, almost surely, nonzero? Here we derive an upper bound on this threshold. Computer experiments that we have conducted suggest that this bound may coincide with the actual threshold, but this remains an open question.

1 Introduction

We study random simplicial complexes in the $X_d(n, p)$ model which was introduced in [7] and further studied in [9, 6, 3, 2, 1]. We quickly recall the basic features of this model. A simplicial complex $X \sim X_d(n, p)$ has n vertices, and a full $(d-1)$-dimensional skeleton. The d-dimensional faces of X are selected uniformly and independently with probability $1 \geq p \geq 0$. The parameter p may (and actually will usually) be dependent on n. We fix once and for all an arbitrary field $\mathbb{F}$ and let $H_i(X) = H_i(X; \mathbb{F})$ denote the i-th homology group of X with coefficients in $\mathbb{F}$, and $h_i(X) = \dim_{\mathbb{F}} H_i(X)$. The

[*]Department of Computer Science, Hebrew University, Jerusalem 91904, Israel. e-mail: liobsar@gmail.com .

[†]Department of Computer Science, Hebrew University, Jerusalem 91904, Israel. e-mail: nati@cs.huji.ac.il . Supported by ISF and BSF grants.

main result of this paper is an upper bound on the threshold for the almost sure nonvanishing of $H_d(X)$ for $X \sim X_d(n, p)$. This question was addressed in [1] and the present paper is, in many ways, a continuation of that paper.

To put this result in perspective, it may be useful to recall the situation for $d = 1$ in which case $X_d(n, p)$ coincides with the Erdős-Rényi $G(n, p)$ model of random graphs. For every $1 > c > 0$, a random graph from $G(n, \frac{c}{n})$ contains a cycle with some probability $1 > f(c) > 0$. Therefore, the threshold for the appearance of a 1-homology in $G(n, \frac{c}{n})$ is coarse. Namely, the probability that the random graph from $G(n, \frac{c}{n})$ contains a cycle, jumps from a positive number to 1 as c changes from $1 - \epsilon$ to 1. The behavior for dimension $d > 1$ is similar. This is due to the fact that $\partial \Delta_{d+1}$, the boundary of a $(d+1)$-simplex occurs with positive probability in $X_d(n, \frac{c}{n})$ for every $c > 0$.

However, the situation in one dimension and in higher dimensions are quite different when we consider the size of the first occurring cycle. This issue is a little easier to discuss in the language of a random graph (or complex) process, where edges (resp. d-faces) are added at random one at a time. For graphs, the length of the first generated cycle is distributed according to a certain known distribution [4]. In contrast, in the random d-dimensional complex process, it follows from [1] that the first emerging cycle in $H_d(X)$ is either $\partial \Delta_{d+1}$, or it has cardinality $\Omega(n^d)$. We have conducted fairly extensive computer experiments with $d = 2$. In these experiments the situation was *always* that in the latter case, the first emerging cycle included *all* n vertices. Whether this can be proved and whether the same holds for general d remains a subject for further study.

The *degree* of a $(d-1)$-face in a d-dimensional complex is the number of d-faces that contain it. A $(d-1)$-face of degree zero, i.e. one that is contained in no d-face is said to be *isolated*. A $(d-1)$-face of degree 1 is said to be *free*. The removal of a free $(d-1)$-face and the unique d-face that contains it is called an *elementary collapse*. We recall [5] that an elementary collapse is a homotopy equivalence. Given a complex X, we carry out a series of elementary collapses that take place in *phases*. At the beginning of a phase we list all $(d-1)$-faces in the complex which are currently free, and we scan them in an arbitrary order. As we arrive at a $(d-1)$-face τ in the list, one of two things can happen. It may still be free, in which case we apply to it an elementary collapse. It is also possible that when τ is reached, it is already isolated, since the unique d-face that initially contained it was already eliminated in a previous elementary collapse. In this case we simply skip τ. When we reach the end of the list, the current phase terminates and

2

a new phase commences.

We denote by $R_i(X)$ the complex gotten from X at the end of phase i. In particular $R_0(X)$ is just the randomly drawn complex with which we start. A d-face from $R_{i-1}(X) \setminus R_i(X)$ is said to be in *generation i*. A $(d-1)$-face is of *generation i* if its degree in $R_{i-1}(X)$ is positive and it either does not belong to $R_i(X)$ or is isolated there.

Theorem 1.1. *Let $1 > \beta = \beta_d > 0$ be the unique positive root of the equation*

$$- \ln(1 - \beta) = \frac{(d+1) \cdot \beta}{d + 1 - d \cdot \beta} \tag{1}$$

and let c_d^* be defined as

$$c_d^* = \frac{- \ln(1 - \beta)}{\beta^d}. \tag{2}$$

If $c > c_d^*$, then a complex X drawn from $X_d(n, \frac{c}{n})$ satisfies asymptotically almost surely

$$H_d(X) \neq 0.$$

Specifically, $c_2 = 2.75381, c_3 = 3.90708$. Also $\beta_2 = 0.883414$ and $\beta_3 = 0.972498$. For large d

$$c_d^* = (d+1) - \frac{d^2 + d + 1}{\exp(d+1)} + O(\frac{d^2}{\exp(2d)})$$

and $\beta_d = 1 - \exp(-(d+1)) - (1 + o_d(1)) \frac{(d+1)^2}{\exp(2(d+1))}$.

Note: Here are a few words about our experiments for $d = 2$. We run the random process in which a random 2-face is added to the complex at each step. The experiment splits according to whether the first cycle to occur is $\partial \Delta_3$ or not. Conditioned on the first cycle not being $\partial \Delta_3$, the numerical estimates that we get for c_2^* for $n = 50, 100, 200$ are (expectation $\pm$ standard deviation) $2.70424 \pm 0.03115, 2.72886 \pm 0.01534, 2.74149 \pm 0.00733$ respectively. This lends some support to our belief that the bounds attained in Theorem 1.1 is the true value of the threshold probability.

The general strategy of our proof is this: An elementary collapse is a homotopy equivalence and in particular it preserves the homology of the complex. Using ideas similar to [1] we observe what happens as we systematically collapse (in phases, as described above) every free $(d-1)$-face. An elementary collapse eliminates exactly one $(d-1)$-face and one d-face.

3

However, it also happens that a non-isolated $(d-1)$-face becomes isolated through collapses on other $(d-1)$-faces. We denote by C_*^1 the (random) set of such faces. Also, let C_*^0 be the set of isolated $(d-1)$ faces in the original X. Finally let $C_* := C_*^0 \cup C_*^1$ and let $\zeta_* := |C_*|$. Observe that if $f_d(X) > f_{d-1}(X) - \zeta_*$, then $H_d(X) \neq 0$. Thus the main parts of the proof are these:

- Local analysis of the collapsing process.

- Computing the expectation $\mathbb{E}(\zeta_*)$.

- A measure-concentration argument on the random variable ζ_*.

Our proof uses the fact that every $(d-1)$-face in C_* corresponds to a zero row in the inclusion matrix of $(d-1)$-faces vs. d-faces (after the collapses). There is another way of establishing the threshold, as done in the upper bound proof in [1]. It is possible to associate to every $(d-1)$-face in C_* a cocycle in $Z^{d-1}(X)$ and use the Euler-Poincaré relation to give an upper bound on the threshold. Indeed our proof can be viewed as an extension of the argument of [1].

As the reader has probably noticed, the above explanation says nothing about the parameter β which plays a key role in the theorem. This is done in Section 4.1 below, where we provide a more comprehensive overview of the proof.

2 The Probability Space $\mathcal{T}_d(k, c)$

We analyze the sequence of d-complexes which are obtained, starting from X and repeatedly collapsing, in phases. Our analysis seeks to determine the way at which a given face ϕ of dimension $(d-1)$ or d gets collapsed. Note that ϕ's generation is completely determined by its local neighborhood in X. Concretely, if ϕ is of generation k, this can be ascertained by observing ϕ's radius-$(k+1)$ neighborhood. For every fixed k this neighborhood is almost surely a d-tree. We analyze the properties of this neighborhood using an intermediary – A Galton-Watson-like model of d-trees. This model is relatively easy to comprehend, and yet it provides a good approximation to the true local behavior of $X_d(n, \frac{c}{n})$ at the vicinity of ϕ. This general strategy has been used numerous times, and in particular in [1].

We turn to provide the necessary definitions. We start with the (recursive) definition of a *d-tree*. A single *d*-face is a *d*-tree. A *d*-tree on $n + 1$ vertices is obtained by taking a *d*-tree T on n vertices and adding to it a new *d*-face $v \cup \tau$ and its $(d-1)$-skeleton. Here τ is a $(d-1)$-face of T, and v is a new vertex. A *rooted d-tree* is a *d*-tree in which we designate one $(d-1)$-face to be the root.

Associated with every *d*-complex Z is a graph G_Z whose vertices are the *d*-faces and the $(d-1)$-faces of Z. An edge between a *d*-face and a $(d-1)$-face stands for inclusion, and two $(d-1)$-faces are adjacent when they are contained in a *d*-face of Z. We freely apply to Z graph-theoretic notions from G_Z such as distance, diameter and radius from a vertex. Thus the distance between two $(d-1)$-faces in Z is the distance of the two corresponding vertices in G_Z.

We define a probability space $\mathcal{T}_d(k, c)$ of *d*-trees of radius $\leq k$ from a $(d-1)$-face τ that is the root of the tree. Thus $\mathcal{T}_d(0, c)$ is just the root $(d-1)$-face τ. For $k > 0$ we sample a *d*-tree from $\mathcal{T}_d(k, c)$ as follows:

- Sample a *d*-tree T from $\mathcal{T}_d(k - 1, c)$.

- For each $(d-1)$-face θ in T at distance $k - 1$ from τ

  - Sample an integer j from the Poisson(c) distribution.
  - Create j new vertices $t_1, \ldots t_j$ and add j new *d*-faces $\theta \cup t_i$ to T for $i = 1, \ldots, j$.

It is useful now to introduce a variation on the notion of the collapsing process. This is a process which we call θ-*collapsing*, where θ is a $(d-1)$-face. This process is identical to the process of collapsing in phases, except that θ *must not be collapsed*, even when it happens to be free. We analyze how a random *d*-tree $T \in \mathcal{T}_d(k + 1, c)$ behaves under the τ-collapsing process where τ is the root of T. It is obvious that after $k + 1$ phases of τ-collapsing, T collapses to τ, but we need to know whether τ becomes isolated in phase $k + 1$ or sooner. To this end we define the event $\mathcal{C}_r(k + 1, d, c)$ that τ belongs to generation earlier than r, where $r \leq k + 1$. We denote the probability of $\mathcal{C}_r(k + 1, d, c)$ by $\gamma_r(k + 1, d, c)$. As mentioned above, whether or not τ becomes isolated at time $< r$ depends only on its radius-r neighborhood in T. In particular, $\gamma_r(k + 1, d, c) = \gamma_r(r, d, c)$ if $k + 1 \geq r$. We denote $\gamma_r(r, d, c)$ by $\gamma_r(d, c)$. Let us calculate the numbers $\gamma_r(d, c)$ for small r. Clearly $\gamma_0(d, c) = 0$.

Also, $\mathcal{C}_1(k+1, d, c)$ is the event where T consists only of its root τ, so that

$$\gamma_1(d, c) = e^{-c}. \tag{3}$$

Notice that τ, the root of a tree becomes isolated before the r-th phase iff each d-face $\sigma \supset \tau$ satisfies the following condition. There is a $(d-1)$-subface $\sigma \supset \tau'$ that we view as the root of a d-tree T' which we τ'-collapse. In the τ'-collapsing of T', the root τ' becomes isolated before phase $(r-1)$. Let π_j be the probability that a Poisson(c) random variable takes the value j we obtain:

$$\gamma_r(d, c) = \sum_{j=0}^{\infty} \pi_j(1 - (1 - \gamma_{r-1}(d, c))^d)^j =$$

$$\sum_{j=0}^{\infty} \frac{c^j}{j!} e^{-c}(1 - (1 - \gamma_{r-1}(d, c))^d)^j = \tag{4}$$

$$\exp(-c(1 - \gamma_{r-1}(d, c))^d) .$$

We denote by $\mathcal{B}_k(d, c)$ the event that the root of $T \in \mathcal{T}_d(k+1, c)$ belongs to a generation later than k. The probability of this event is $\beta_k(d, c)$. Clearly

$$\beta_k(d, c) = 1 - \gamma_{k+1}(d, c). \tag{5}$$

3   The Neighborhood of a $(d-1)$-face

The next step is rather standard in arguments of the sort we are making. Most of the necessary details are to be found in [1], and we now provide a few additional comments and explanations. The purpose is to show that the Poisson-distribution-based tree considered above approximates arbitrarily closely (as $n \to \infty$) the actual local behavior of our random complex.

How does the neighborhood of a $(d-1)$-face τ in a d-dimensional complex X look like? The 0-neighborhood $\mathcal{S}_0$, consists of τ alone. The i-th neighborhood $\mathcal{S}_i$ is the complex generated by the d-faces in $\mathcal{S}_{i-1}$, and the additional d-faces that contain a $(d-1)$-face in $\mathcal{S}_{i-1}$. We denote by v_i the number of vertices in $\mathcal{S}_i$.

Let A_k be the event (in $X_d(n, p)$) that $\mathcal{S}_k$ is a d-tree. Let D be the event that every $(d-1)$-face in $X \sim X_d(n, p)$ has degree $\leq \log n$.

The argument in [1] has two parts. One shows first

Claim 3.1. *Let k and $c > 0$ be fixed and $p = \frac{c}{n}$. Then*

$$\Pr[A_k \cap D] = 1 - o(1).$$

The next step is to show that conditioned on the event $A_k \cap D$ the following recursive random process generates the typical k-neighborhood in X. As before $\mathcal{S}_0 = \tau$ and for $i > 0$, with $\mathcal{S}_{i-1}$ already in place, the next layer $\mathcal{S}_i$ is generated according to the following rule: For each $(d-1)$-face θ in $\mathcal{S}_{i-1}$ at distance $i - 1$ from τ

- Sample an integer j from the binomial distribution $B(n - v_{i-1}, \frac{c}{n})$.

- Create j new vertices $t_1, \dots t_j$ and add j new d-faces $\theta \cup t_i$ to $\mathcal{S}_i$ for $i = 1, \dots, j$.

The only difference between this random process and the way we defined $\mathcal{T}_d(k, c)$ is that we sample the integer j from $B(n - v_{i-1}, \frac{c}{n})$ and not from Poisson(c). Notice that if $X \in A_k \cap D$ then $v_k = O(\log^k n)$ and the total variation distance between $\mathcal{S}_k$ and $\mathcal{T}_d(k, c)$ is $o(1)$.

Thus for a $(d-1)$-face τ

- $\Pr(\deg_{R_k(X)}(\tau) > 0) = (1 - o(1))\beta_k(d, c)$,

- $\Pr(\deg_{R_{k-1}(X)}(\tau) = 0) = (1 - o(1))\gamma_k(d, c))$.

Consider an inclusion $\tau \subset \sigma$ of a $(d-1)$-face and a d-face. Let S'_k be the k-th neighborhood of τ in $X \setminus \sigma$. We can apply Claim 3.1 to S'_k and conclude that with probability $1 - o(1)$ it is a d-tree in which every $(d-1)$-face has degree at most $\log n$. To randomly generate S'_k we just run the random process that generates S_k and modify it, by excluding σ from S'_1. Thus

- $\Pr(\deg_{R_k(X) \setminus \sigma}(\tau) > 0) = (1 - o(1))\beta_k(d, c)$,

- $\Pr(\deg_{R_{k-1}(X) \setminus \sigma}(\tau) = 0) = (1 - o(1))\gamma_k(d, c))$.

4 An Upper Bound For The Threshold

As usual, we associate a boundary operator with the d-dimensional complex X. This linear operator corresponds to an $f_{d-1}(X) \times f_d(X)$ matrix M whose rows and columns are indexed by X's $(d-1)$ resp. d-faces. All entries of M

are in $\{-1, 0, 1\}$ and are defined as follows. Every d-face $\sigma \in X$ is given an orientation $[v_0, \ldots, v_d]$, and $M_{\sigma \setminus v_i, \sigma} := (-1)^i$ for every $0 \leq i \leq d$. All other entries of M equal 0. Since X is d-dimensional, $h_d(X) = 0$ iff M's right kernel is zero.

Let N be an $a \times b$ matrix and let $\zeta = \zeta(N)$ be the number of zero rows in N. Clearly N has a nonzero right kernel if $b > a + \zeta$. We apply this simple observation to M_i, the matrix associated with $R_i(X)$. Since an elementary collapse is a homotopy equivalence, $H_d(R_i(X)) = H_d(X)$ for all i. We conclude that if $s_i(X) \overset{\text{def}}{=} f_d(R_i(X)) - f_{d-1}(R_i(X)) + \zeta_i(X) > 0$ for some $i \geq 0$, then $H_d(X) \neq 0$.

Our proof shows that if $c > c_d^*$ then $X \sim X_d(n, \frac{c}{n})$ satisfies a.s $s(X) = s_{k_*}(X) > 0$. Here k_* is a large enough constant to be determined later. First we calculate the expectation of $s(X)$, and then show that a.s $s(X) > 0$.

Theorem 4.1. *Let $p = \frac{c}{n}$ with $c > c_d^*$. Then*

$$\Pr\,[\,X \in X_d(n, p) : s(X) > 0] = 1 - o_n(1)\,.$$

As the previous discussion shows, this theorem implies Theorem 1.1.

Proof: Fix c and d and let β_k, γ_k stand for $\beta_k(d, c), \gamma_k(d, c)$ resp. and let us fix an arbitrarily small $\epsilon > 0$. Let $\zeta_*(X) = \zeta(M_{k_*}(X))$.

A d-face σ is in $R_{k_*}(X)$ if $\sigma \in X$ and it is not collapsed in phase k_* or earlier. In particular, at the end of phase $k_* - 1$, every $(d-1)$-subface of σ must be contained as well in a d-face other than σ. Hence

$$\mathbb{E}[f_d(R_{k_*}(X))] = (1 - o(1))\binom{n}{d+1}\frac{c}{n}\beta_{k_*-1}^{d+1} \qquad (6)$$

Let τ be a $(d-1)$-face. Let $\mathcal{Q}_\tau$ be the event that τ becomes isolated after k_* τ-collapsing phases. The discussion in Sections 2 and 3 yields that $\Pr(\mathcal{Q}_\tau) = (1 - o(1))\gamma_{k_*+1}$. Let $\mathcal{P}_\tau$ be the event that τ collapses after k_* collapsing phases, but does not become isolated after k_* τ-collapsing phases. Equivalently, τ becomes a free subface of some d-face σ before collapsing phase k_*, but all other $(d-1)$-subfaces of σ are not free prior to collapsing phase k_*. Consequently, $\Pr(\mathcal{C}_\tau) = (1 - o(1))\frac{c}{n}n\gamma_{k_*}\beta_{k_*-1}^d$.

The row corresponding to a $(d-1)$-face τ becomes a zero row or is removed from the matrix after the collapsing phases if τ is either isolated or was collapsed. Hence this happens only if the event $\mathcal{Q}_\tau \cup \mathcal{P}_\tau$ occurs. Thus the

probability that some $(d-1)$-face belongs to the complex and is not isolated after k_* collapsing phases is $(1-o(1))(1-(\gamma_{k_*+1}+c\cdot\gamma_{k_*}\beta_{k_*-1}^d))$. Therefore

$$\mathbb{E}[f_{d-1}(R_{k_*}(X))-\zeta_{k_*}(X)] = (1-o(1))\binom{n}{d}(1-(\gamma_{k_*+1}+c\cdot\gamma_{k_*}\beta_{k_*-1}^d)) \quad (7)$$

Consequently,

$$\mathbb{E}[s] = \mathbb{E}[f_d(R_{k_*}(X))] - \mathbb{E}[f_{d-1}(R_{k_*}(X))+\zeta_*]$$

$$= (1-o(1))\binom{n}{d}\left(-\beta_{k_*}+c\beta_{k_*-1}^d(1-\beta_{k_*-1})+c\frac{\beta_{k_*-1}^{d+1}}{d+1}\right). \quad (8)$$

4.1 overview

We are now ready to provide a more detailed explanation of our strategy of proof for Theorem 1.1. We fix an integer $d \geq 2$ once and for all, and some $c > 0$ whose value we discuss below. With this fixed c, we obtain a recurrence relation for β_k. (This follows readily from Equation (4)). Namely, starting with $t = 1$ we recurse on $t \to 1 - f(t)$, where $f_c(t) = f(t) = \exp(-c\cdot t^d)$. It is easily verified that for every $c > 0$ the function $1 - f(\cdot)$ is increasing in $[0, 1]$. As already observed in [1] there is some $c_{\text{collapse}} > 0$ depending only on d so that when $c_{\text{collapse}} > c > 0$, the only root of $1 - f_c(t) = t$ in $[0, 1]$ is $t = 0$. Therefore, for $c_{\text{collapse}} > c > 0$ the recurrence $t \to 1 - f(t)$ started at $t = 1$ converges to zero.

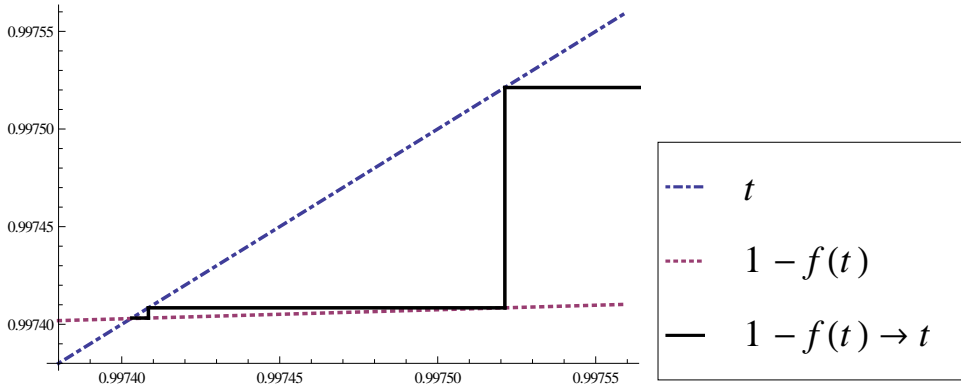Figure 1: The recurrence relation for β_k

9

Routine calculations (see also Figure 2) yield that for $c > c_{\text{collapse}}$ there are exactly two roots in $(0, 1)$ to $1 - f_c(t) = t$. In this range the above recurrence converges to the larger of these two roots. As Equation (8) shows, $\mathbb{E}(s) > 0$ iff

$$- \ln(1 - \beta) > \frac{(d+1) \cdot \beta}{d + 1 - d \cdot \beta}$$

In the statement of Theorem 1.1 the same calculations are done "in reverse". Namely, Equation (2) states that $1 > \beta > 0$ is a root of $1 - f_{c_d^*}(t) = t$. It only remains to rule out the possibility that Equation (1) yields the *smaller* root. To this end, note that the solution of Equation (1) is $\beta > 1 - \exp(-d)$. Moreover, the larger/smaller root of $1 - f_c(t) = t$ increases resp. decreases with c and the smaller root is smaller than $1 - \exp(-d)$ for every $c > c_{\text{collapse}}$.
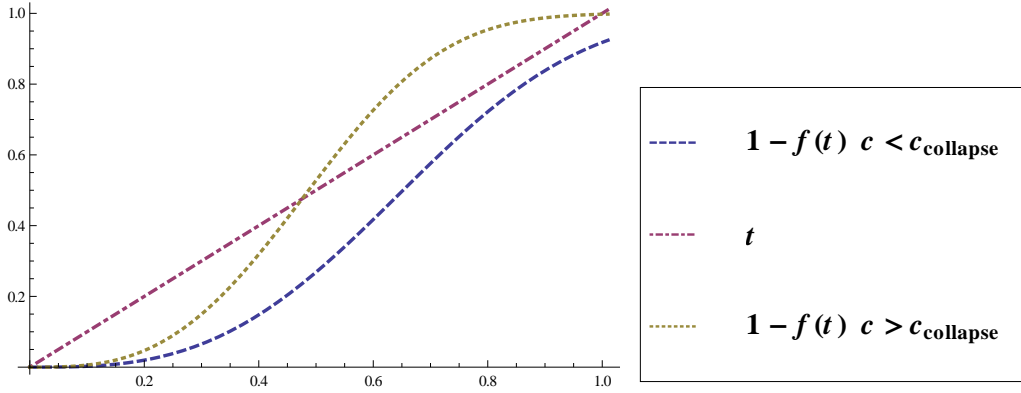

Figure 2

Let k_* be large enough s.t $\beta_{k_* - 1} - \beta_* < \epsilon$. Since $c > c_d^*$ we conclude that for n large enough

$$\mathbb{E}[s] \geq \epsilon' \binom{n}{d} \tag{9}$$

where $\epsilon' > 0$ depends only on c and d.

4.2 A concentration of measure argument

The only missing part of the proof is that $s > 0$ almost surely. This is shown using the following version of Azuma's inequality from [8].

Theorem 4.2. *Let $Y_1, \ldots, Y_m$ be random variables taking values from $\{0, 1\}$. Suppose $\Phi : \{0, 1\}^m \to \mathbb{R}$ satisfies:*

$$|\Phi(x) - \Phi(x')| \leq \epsilon_k$$

for x and x' that differ only at their k-th coordinate. Then for every $t > 0$:

$$\Pr[|\Phi(Y_1, \ldots, Y_m) - \mathbb{E}[\Phi(Y_1, \ldots, Y_m)]| \geq t] \leq 2e^{-\frac{2t^2}{\sum_k \epsilon_k^2}} \qquad (10)$$

Let $\sigma_1, \ldots, \sigma_{\binom{n}{d+1}}$ be the list of all d-faces, and let Y_i be the indicator random variable of the event $\sigma_i \in X$. We apply this theorem with $m = \binom{n}{d+1}$ and with $\Phi = s$. Let us consider two d-complexes X and X' that are identical, except that $\sigma \in X$, but $\sigma \notin X'$. We need to provide an upper bound on $|s(X) - s(X')|$. In an elementary collapse we eliminate one d-face and one $(d-1)$-face. Thus, in particular $f_d(R_i(X)) - f_{d-1}(R_i(X)) = f_d(X) - f_{d-1}(X)$ for every i. Clearly $f_d(X) - f_d(X') = 1$ and $f_{d-1}(X) = f_{d-1}(X')$, so we only need a bound on $|\zeta_*(X') - \zeta_*(X)|$. As we show

$$(d+1) \cdot d^{k_*} \geq |\zeta_*(X') - \zeta_*(X)|.$$

We now compare the collapsing processes as they evolve in X and in X'. If a $(d-1)$-face is of generation $i, i' \leq k_*$ in X, X' respectively, then clearly $i \geq i'$. Let Θ be the set of $(d-1)$-faces in X for which $i > i'$. Clearly Θ is contained in the k_*-neighborhood of σ. We classify the faces in Θ according to their distance from σ and show that at distance j from σ there are at most $(d+1) \cdot d^{j-1}$ members of Θ. This is clearly true for $j = 1$, namely the $d+1$ subfaces of σ. The general claim is shown by induction on j. A $(d-1)$-face $\tau \in \Theta$ at distance $j+1$ from σ must have a neighbor (in G_X), a $(d-1)$-face $\tau' \in \Theta$ whose distance from σ is j so that τ' is of generation one earlier than τ. In particular, there is a d-face φ that contains both these τ and τ' and is collapsed through τ'.

So let $\tau' \in \Theta$ be a face of generation ν' in X' whose distance from σ is j. Since τ' can collapse only one d-face, it can have at most d neighbors in Θ, which are $(d-1)$-faces at distance $j+1$ from σ. We can conclude that there are at most $(d+1)d^j$ faces in Θ at distance $j+1$ from σ.

Clearly $|\zeta_*(X') - \zeta_*(X)| \leq |\Theta|$, and $|\Theta| \leq \sum_{j=1}^{k_*}(d+1)d^{j-1} \leq (d+1)d^{k_*}$ by the previous discussion. Thus $|s(X) - s(X')| \leq (d+1)^{k_*+1}$, by using (9) and (10) we conclude:

11

$$\Pr[s < 0] \leq 2e^{-\frac{2\mathbb{E}^2[s]}{\binom{n}{d+1}(d+1)^{2k_*+2}}} = o(1).$$

$\square$

5 Open problems

- The most obvious remaining challenge is to determine the correct threshold for the non-vanishing of the d-th homology in $X_d(n, p)$. As stated before we believe c_d^* is this threshold.

- The present results and those of [1] strongly suggest that the threshold for collapsibility is substantially smaller than the one for the almost sure non-vanishing of the d-th homology. Can one at least show that the two thresholds do not coincide?

- In the random complex process, what is the distribution of the first emerging cycle in $H_d(X)$? In particular, can one prove that (as suggested by our numerical experiments) it is either $\partial\Delta_{d+1}$ or else it includes all n vertices?

- It would be extremely interesting to investigate the inclusion matrices of $(d-1)$-faces vs. d-faces of complexes in $X_d(n, \frac{c}{n})$ for values of c between the two thresholds (assuming, of course, that they differ). If the conjectures alluded to in the above questions hold, then this matrix has excellent properties, when viewed as the parity-check matrix of an error-correcting code.

6 Acknowledgment

References

[1] L. Aronshtam, N. Linial, T. Łuczak and R. Meshulam, Collapsibility and vanishing of top homology in random simplicial complexes, arXiv:1010.1400

[2] E. Babson, C. Hoffman and M. Kahle, The fundamental group of random 2-complexes, *J. Am. Math. Soc.* 24(1), 128 (2011)

[3] A. Costa, M. Farber and T. Kappeler, Topology of random 2-complexes, arXiv:1006.4229 .

[4] P. Flajolet, D. E. Knuth, and B. Pittel, The first cycles in an evolving graph, *Discrete Math*, 75(1989), 167-215 .

[5] A. Hatcher, Algebraic Topology, Cambridge University Press, 2002.

[6] D. Kozlov, The threshold function for vanishing of the top homology group of random d-complexes, *Proc. Amer. Math. Soc.*138(2010) 45174527

[7] N. Linial and R. Meshulam, Homological connectivity of random 2-dimensional complexes,*Combinatorica*, 26(2006) 475–487.

[8] C. McDiarmid, On the Method of Bounded Differences, *Surveys in combinatorics 1989*, 148–188, London Math. Soc. Lecture Note Ser., 141, Cambridge Univ. Press, Cambridge, 1989.

[9] R. Meshulam and N. Wallach, Homological connectivity of random k-dimensional complexes, *Random Struct. Algorithms*, 34(2009) 408-417